

Die Virtuelle Fachbibliothek als Schnittstelle für eine fachübergreifende Suche in den einzelnen Virtuellen Fachbibliotheken

Definition eines Metadaten-Kernsets (VLib Application Profile)

Hans Jürgen Becker, Christel Hengel, Heike Neuroth,
Berthold Weiß, Carola Wessel¹

1. Vorbemerkung und Einführung

Die Deutsche Forschungsgemeinschaft (DFG) hat in ihrem 1998 verfassten Memorandum zur Weiterentwicklung des Systems der überregionalen Literaturversorgung² die Bildung Virtueller Fachbibliotheken angeregt. Über den Aufbau der Virtuellen Fachbibliotheken in der Bundesrepublik Deutschland wurde im BIBLIOTHEKSDIENST in mehreren Beiträgen³ berichtet. Weitere Artikel sind geplant.⁴

-
- 1 An der Erstellung der Empfehlungen haben auch *Jürgen Braun*, SUB Göttingen und *Susanne Nickel*, Die Deutsche Bibliothek mitgewirkt.
 - 2 Deutsche Forschungsgemeinschaft: Memorandum zur Weiterentwicklung der überregionalen Literaturversorgung. In: ZfBB 45 (1998), S. 135-164; online unter: <<http://www.dfg.de/foerder/biblio/memo.html>>.
 - 3 Die VLib Partner haben beschlossen, in einer Reihe von Publikationen ihre Virtuellen Fachbibliotheken im BIBLIOTHEKSDIENST vorzustellen. Als erster Beitrag erschien der Übersichtsartikel: Sven Meyenburg: Der Aufbau Virtueller Fachbibliotheken in der Bundesrepublik Deutschland. In: BIBLIOTHEKSDIENST 34 (2000), H.7/8, S. 1229-1235. Danach folgten: Wolfgang Meier, Matthias N.O. Müller, Stefan Winkler: Virtuelle Fachbibliothek Sozialwissenschaften. In: BIBLIOTHEKSDIENST 34 (2000), H. 7/8, S. 1236-1244; Ulrich Tiedau, Peter te Boekhorst: NedGuide – Virtuelle Fachbibliothek Niederländischer Kulturkreis. In: BIBLIOTHEKSDIENST 34 (2000), H. 10, S. 1687-1694; Reimer Eck, Alexander Huber: Die Virtuelle Fachbibliothek Anglo-Amerikanischer Kulturkreis, Fachgebiet Anglistik (=VLib-AAC Lit) bei der SUB Göttingen. In: BIBLIOTHEKSDIENST 35 (2001), H. 1, S. 38-48; Gudrun Gersmann; Marianne Dörr: Der Server Frühe Neuzeit als Baustein für eine Virtuelle Fachbibliothek Geschichte. In: BIBLIOTHEKSDIENST 35 (2001), H. 3, S. 283-294; Elzbieta Gabrys-Deutscher: ViFaTec – Die Virtuelle Fachbibliothek Technik. In: BIBLIOTHEKSDIENST 35 (2001), H. 5, S. 573-582; Natascha Magyar: Die Virtuelle Fachbibliothek Psychologie an der Saarländischen Universitäts- und Landesbibliothek. In: BIBLIOTHEKSDIENST 35 (2001), H. 7/8, S. 853-867; Lutz Wiederhold: MENALIB – Die Virtuelle Fachbibliothek Vorderer Orient / Nordafrika. In: BIBLIOTHEKSDIENST 35 (2001), H. 10, S. 1327-1336. Vgl. auch <<http://bibliotheksdienst.zib.de/>>

Der folgende Beitrag beschäftigt sich nicht mit einer konkreten Virtuellen Fachbibliothek, sondern mit dem übergreifenden Themenkomplex „Metadaten“ und der Fragestellung, wie Metadaten für eine fachübergreifende Suche über alle Virtuellen Fachbibliotheken eingesetzt werden können.

Im Rahmen des Aufbaus der Virtuellen Fachbibliotheken hat die Projektkoordinierung Unterarbeitsgruppen zur Lösung spezifischer Fragestellungen ins Leben gerufen. Der Arbeitsbereich „Metadaten“ ist von dem von der DFG geförderten Projekt META-LIB (Metadaten-Initiative deutscher Bibliotheken⁵) mit den Teilprojekten an Der Deutschen Bibliothek⁶ und der SUB Göttingen⁷ übernommen worden. META-LIB erhielt die Aufgabe, „Empfehlungen zur Definition eines Metadaten-Core-Sets für die verteilte Suche über die Virtuellen Fachbibliotheken“ zu entwickeln.

Im folgenden werden die Empfehlungen vorgestellt. Sie basieren auf den Ergebnissen und der Auswertung von Antworten⁸ eines Internet-Fragebogens⁹, in dem die Datenelemente erfragt wurden, die in den einzelnen Virtuellen Fachbibliotheken zur Erschließung verwendet bzw. benötigt werden. Für die Formulierung der Empfehlungen und zur Abstimmung sind zwei Metadaten-Workshops (am 16. Mai 2001 in der SUB Göttingen und am 9./10. August 2001 in der Deutschen Bibliothek Frankfurt am Main) durchgeführt worden, deren Ergebnisse und Diskussionen als Grundlage dienen. Darüber hinaus sind Erfahrungen, die bei der Entwicklung des Renardus¹⁰ Application Profiles¹¹ gesammelt wurden, hier eingeflossen. Generell empfiehlt es sich, bei der Erstellung eines Application Profiles internationale Entwicklungen und Diskus-

-
- 4 Für das kommende Heft ist ein Beitrag über die Virtuelle Fachbibliothek Politikwissenschaft und Friedensforschung geplant.
 - 5 Siehe dazu den aktuellen Artikel: Carola Wessel; Berthold Weiß: META-LIB. Die Metadaten-Initiative deutscher Bibliotheken. In: Bibliothek. Forschung und Praxis 25 (2001), Nr. 3, S. 301-305; online: <http://www.bibliothek-saur.de/2001_3/301-305.pdf>
 - 6 <<http://www.ddb.de/professionell/metalib.htm>>
 - 7 <http://www2.sub.uni-goettingen.de/sub_de.html>
 - 8 VLib Antworten: <http://www.ddb.de/professionell/metalib_vlib.htm> oder <<http://www2.sub.uni-goettingen.de/metacore/antworten/>> bzw. auch Auswertung <<http://www2.sub.uni-goettingen.de/metacore/ergebnisse/>>
 - 9 VLib Fragebogen: <<http://www.ddb.de/etc/FormularTeil1.html>> oder <<http://www2.sub.uni-goettingen.de/metacore/fragebogen/>>
 - 10 Renardus Projekt <<http://www.renardus.org/>>
 - 11 Renardus Application Profile – Overview <<http://renardus.sub.uni-goettingen.de/renap/>>

sionen zu berücksichtigen, um einerseits die Interoperabilität mit anderen Anwendungen (z.B. zwecks Metadatenaustausch) und andererseits die Nutzbarkeit dieser Application Profiles zu erhöhen (z.B. durch Registries, Suchmaschinen etc.).

Beim Aufbau Virtueller Fachbibliotheken werden von den einzelnen Institutionen unterschiedliche Vorgehensweisen verfolgt. Für den Nachweis von Netzpublikationen werden vielfach vom herkömmlichen Katalog getrennte Nachweissysteme aufgebaut. Ziel der vorgelegten Empfehlungen ist es, eine gemeinsame Suche über verteilt vorkommende, heterogene Datenbestände zu ermöglichen. Es werden demnach Metadaten für Publikationen aller Medienformen aus verteilt liegenden Informationssystemen in die gemeinsame Suche einbezogen, um für den Benutzer eine „gebündelte Informationsversorgung aus einer Hand möglichst ohne Medienbruch“¹² zu ermöglichen. Dafür ist eine „Normalisierung“ des lokalen Metadatenformats der einzelnen Virtuellen Fachbibliothek auf das gemeinsame Metadaten-Set (VLib Application Profile) notwendig.

Als international anerkannter Standard bietet sich das Dublin Core Metadata Element Set¹³ mit Qualifiern¹⁴ an, welches schon vielfach in den einzelnen Virtuellen Fachbibliotheken zur Beschreibung von Netzpublikationen verwendet wird. Die Transformation von Katalog-Metadaten erfolgt über bestehende Konkordanzanzen zwischen den verwendeten Formaten (MAB, MARC und PICA) und Dublin Core.¹⁵

Die vorgelegten Empfehlungen gelten im Prinzip für alle Ressourcen-Typen, orientieren sich zunächst aber an Textressourcen. In einzelnen Modulen der Virtuellen Fachbibliotheken werden aber zum Beispiel auch Multimedia-Ressourcen, Events und Software-Programme nachgewiesen. Das empfohlene Set muss für die Beschreibung dieser Dokumenttypen noch erweitert werden.

12 Rutz: SSG-Programm, Virtuelle Fachbibliotheken und das Förderkonzept der DFG. In: Bibliothek 22 (1998), 3, S. 303-308, hier S. 305.

13 Dublin Core Metadata Element Set, Version 1.1: Reference Description
<<http://www.dublincore.org/documents/dces/>>

14 Dublin Core Qualifiers <<http://www.dublincore.org/documents/dcmes-qualifiers/>>

15 Zur Dokumentation von Konkordanzanzen zwischen Dublin Core und MAB, siehe:
<http://www.ddb.de/professionell/dc_mab.htm>; Dublin Core und PICA:
<<http://www2.sub.uni-goettingen.de/metaform/crosswalks.html#mappings>>; Dublin Core und MARC: <<http://lcweb.loc.gov/marc/dccross.html>>;
<<http://lcweb.loc.gov/marc/marc2dc.html>>

Die Empfehlungen haben neben der gemeinsamen Suche auch das Ziel, Interoperabilität zwischen den einzelnen Virtuellen Fachbibliotheken sowie mit anderen, nationalen und internationalen Kooperationspartnern zu ermöglichen. Dementsprechend sollten die vereinbarten Kernelemente von jeder einzelnen Fachbibliothek verwendet werden. Auf diese Weise kann ein direkter Metadaten-Austausch zwischen zusammenarbeitenden Virtuellen Fachbibliotheken erfolgen. Ergänzend zu dem gemeinsamen Metadaten-Kernset kann die einzelne Virtuelle Fachbibliothek zusätzlich weitere Metadatenelemente definieren bzw. weitere Differenzierungen (Refinements, Encoding Schemes) vornehmen, um ihre lokalen oder fachspezifisch bedingten Anforderungen abbilden zu können. Um dabei die gegenseitige Interoperabilität zu wahren, sollten die in der Dublin Core Metadata Initiative (DCMI) vereinbarten Grundprinzipien eingehalten werden: Neue, qualifizierte Element müssen unter Beachtung des „Dumb-Down-Prinzip“ gebildet werden. Dieses Prinzip besagt, dass bei dem Wegfall der Qualifizierung (Element Refinement oder Encoding Scheme) bei Elementen die inhaltliche Definition der Metadatenelemente (Semantik) gewahrt bleibt, auch wenn die Inhalte an Spezifizierung wie zum Beispiel die Angabe von Title.Alternative bei dem DC Element Title verlieren.

Bei Erweiterungen, d.h. bei Angaben von zusätzlichen Elementen zum Core-Set, ist darauf zu achten, dass das Core-Set in jedem Fall erfüllt bleibt und Inhalte, die in den Core-Elementen untergebracht werden können, in jedem Fall darüber definiert werden. Es wird vorgeschlagen, auch die lokalen Elemente (einschließlich administrative Metadaten) zentral zu dokumentieren, so dass auch andere diese nachnutzen können.

2. Application Profile und Datenmodell

Bei Metadatenformaten ist zu unterscheiden zwischen „Namespace“ und „Application Profile“. Unter einem „Namespace“ wird ein definiertes Set an Metadaten-Elementen verstanden, wie es zum Beispiel im Dublin Core Metadata Element Set (**dc1.1**) mit den 15 fest definierten Elementen vorliegt.¹⁶ Ein weiteres Namespace ist **dcterms**, ebenfalls mit festgelegten Definitionen für einzelne Qualifier¹⁷ und für Metadatenelemente, die nicht in der Dublin Core Ver-

16 <<http://www.dublincore.org/documents/dces/>> Die Bezeichnung dc1.1 ist in den aktuellen Dokumenten noch nicht umgesetzt, vgl. auch *Namespace Policy for the Dublin Core Metadata Initiative (DCMI)*

<<http://www.dublincore.org/documents/2001/10/26/dcmi-namespace/>>

17 <<http://www.dublincore.org/documents/dcmes-qualifiers/>> Hier finden sich auch die Nachweise zu den einzelnen Encoding Schemes. Auch hier ist die Bezeichnung dcterms noch nicht umgesetzt, genauso wie die Integration neuer DC Metadaten-

sion dc1.1 verzeichnet sind. Zu jedem Metadatenelement gibt es neben der semantischen Definition auch syntaktische Beschreibungen, Festlegungen von Wiederholbarkeit und Verpflichtungsgrad sowie „best practice“ Empfehlungen.

Ein solches Set ist allerdings häufig nicht ausreichend für die Bedürfnisse einer bestimmten Anwendergruppe. Anstatt ein gänzlich neues Set zu entwickeln, ist es effizienter, sich aus bereits vorhandenen Namespaces ein für diese Gruppe spezifisches Application Profile zusammenzustellen. Ein Profile besteht demnach aus mehreren Metadatenelementen, die in mehreren verschiedenen Namespaces definiert sein können. Für die spezielle Anwendung werden diese Namespaces zu einem spezifischen Application Profile kombiniert. Dabei ist es möglich, die Definitionen der aus anderen Namespaces übernommenen Elemente zu spezifizieren, wobei diese Definitionen nur spezieller oder semantisch schmäler gefasst, nicht jedoch ausgeweitet werden dürfen¹⁸. Werden über die in vorhandenen Namespaces verwendeten Elemente hinaus weitere spezifische Elemente benötigt, kann man einen eigenen Namespace entwickeln, indem eigene Elemente oder Qualifier definiert werden (vgl. Kapitel Ausblick).

Im Application Profile der Virtuellen Fachbibliothek wird beschrieben, welche Elemente herangezogen werden, welchen Namespaces diese angehören, bei welchen Organisationen die Namespaces registriert, definiert und beschrieben sind und welchen Bedingungen die einzelnen Datenelemente bei der Anwendung in der Virtuellen Fachbibliothek genügen sollen. Außerdem wird eine Praxisempfehlung („best-practice Guide“) für das jeweilige Element und die Elementinhalte gegeben.

Das für die Virtuelle Fachbibliothek empfohlene Application Profile besteht zur Zeit aus fünf Namespaces: Dublin Core Simple (**dc1.1**), Dublin Core Qualified (**dcterms**), der DCMI Type Vocabulary Liste (**dcmitype**) sowie einem einfachen (Virtual Library Metadata Element Set = **VMES**) und einem qualifizierten Set (Virtual Library Metadata Element Set Qualifiers = **VMESq**), das die domainspezifischen Elemente der Virtuellen Fachbibliothek definiert.

elemente (siehe aber *Open Metadata Registry*
<<http://wip.dublincore.org:8080/registry/Registry>>, dort unter „Show all registered: Schemas“).

18 Heery und Patel definieren Application Profiles als “schemas consisting of elements drawn from one or more namespaces, combined together and optimised for a particular application”, Rachel Heery und Manjula Patel: Application profiles: *Mixing and matching metadata schemas* <<http://www.ariadne.ac.uk/issue25/app-profiles/>>. Vgl. auch Makx Dekkers: *Application profiles, or how to mix and match metadata schemas* <<http://www.cultivate-int.org/issue3/schemas>>.

Die folgende Tabelle gibt einen Überblick, in welcher Form das zukünftige VLib Datenmodell definiert bzw. beschrieben wird. Daraus können dann die für die Entwicklung des VLib Application Profiles nötigen einzelnen Namespaces abgeleitet werden.

Name	Name des Metadatenelements
Namespace	Angabe des jeweiligen Namespaces: dc1.1 dcterms dcmitype VMES (VLib Metadata Element Set) VMESq (VLib Metadata Element Set Qualified) In Zukunft sind weitere denkbar, z.B. METADISS (s.u. bei DC.Date).
DC Refinement(s)	DC Element Refinements, die in der Virtuellen Fachbibliothek benutzt werden: Diese Art der Qualifier spezifizieren die Bedeutung des Elements weiter. Ein „Element Refinement“ deckt nur einen Ausschnitt der inhaltlichen Definition (Semantik) des unqualifizierten Elements ab und schränkt es damit ein.
VLib Refinement(s)	VLib Refinement, vgl. oben
DC Encoding Scheme(s)	DC Encoding Scheme benennt value encoding schemes, die die richtige Interpretation des Elementinhaltes ermöglichen. Sie enthalten kontrolliertes Vokabular, Notationen, Syntax-Regeln etc. Ein Wert, der in einem „Encoding Scheme“ ausgedrückt wird, kann entweder eine Zeichenkette (token) aus einem kontrolliertem Vokabular (z.B. ein Term aus einem Klassifikationssystem) oder ein String sein, der entsprechend einer formalen Notation gebildet wurde (z.B. "2000-01-01" als Standardausdruck für eine Datumsangabe). Ist ein value encoding scheme nicht interpretierbar für eine Maschine (z.B. automatische Suche), so kann er für einen menschlichen Leser dennoch nützlich im Sinne einer eindeutigen Interpretierbarkeit sein.
VLib Encoding Scheme(s)	VLib Encoding Scheme, vgl. oben
Grad der Verpflichtung	Es werden vier Grade der Verpflichtung für ein Element benannt: <ul style="list-style-type: none"> • obligatorisch 1: das Element muss immer mit einem Wert versehen sein, • obligatorisch 2: das Element muss, soweit es ermittelt werden kann, immer mit einem Wert versehen sein, • empfohlen (nach Möglichkeit mit einem Wert versehen) • optional (Verwendung freigestellt)
Wiederholbarkeit	Angabe, ob das Element wiederholt werden darf oder nicht

Attribut "LANG"	Das Attribut "LANG": gibt an, in welcher Sprache der Inhalt des jeweiligen Metadatenelements vorliegt. Es wird empfohlen, dass die Sprachangabe nach ISO 639-2 erfolgt. Die Verwendung des Attributs „LANG“ ist in den Feldern empfohlen, bei denen es im Folgenden explizit festgelegt wird.
Attribut „Source of Content“	Dieses Attribut darf nur bei DC.Title und DC.Description angewendet werden. Es gibt an, ob der Feldinhalt vom Autor oder vom redaktionellen Bearbeiter stammt. Als Feldinhalte sind nur die Terme „Autor“, „Third Party“ und „Redaktion“ zulässig. Das Feld soll im gemeinsamen Portal abfragbar sein. Es ist wiederholbar und besitzt den Verpflichtungsgrad „empfohlen“.
DC Definition	Dublin Core Definition des Metadatenelementes
DC Anmerkung	Dublin Core Anmerkungen zu dem Metadatenelement
VLib Definition	VLib Definition des Metadatenelementes
VLib Anmerkung	VLib Anmerkungen zu dem Metadatenelement
Praxisempfehlung	Angaben von Hilfen und Empfehlungen zum Ausfüllen des Inhalts des Metadatenfeldes („best-practice guide“)

Zusätzlich wurde vereinbart, dass bei der Besetzung wiederholbarer Elemente je Element nur ein Wert zulässig ist. Falls mehrere Werte vorliegen, wird das Element entsprechend häufig wiederholt. Zur Beschreibung von Ressourcen, die in mehreren Exemplaren mit unterschiedlichen Adressen (URLs) vorliegen, werden Elementgruppen gebildet, die in einem Metadatensatz abgebildet werden. Die nur für ein bestimmtes Exemplar zutreffenden Metadatenelemente (zum Beispiel: Angaben zum Format, Datum oder zu den Rechten) werden der jeweiligen URL als Elementgruppe zugeordnet.

Um größtmögliche Interoperabilität sowie eine einheitliche Anzeige bei der Suche im Benutzerinterface zu erlangen, soll bei den Eintragungen der Elementwerte möglichst auf schon bestehende Standards zurückgegriffen werden. Diese werden im jeweiligen Metadatenelement als „encoding scheme“ angegeben. Folgende übergreifenden Standards wurden empfohlen: Verwendung von **ISO 3166-1**¹⁹ (**2-Buchstaben-Code**) als **Ländercode**, Verwendung von **ISO 639-2**²⁰ (**3-Buchstaben-Code**) als **Sprachencode**, Verwendung des

19 *ISO 3166 Maintenance Agency (ISO 3166/MA)*
<<http://www.din.de/gremien/nas/nabd/iso3166ma/>>. ISO 3166-1 ist gleichzeitig DIN- und EN-Code. Die Angabe als DIN-Code lautet DIN EN ISO 3166-1.

20 *ISO 639-2 Registration Authority – Library of Congress*
<<http://www.loc.gov/standards/iso639-2/>>

IMT²¹ Standards für die Angabe des Formats der Ressource, Verwendung von **W3C-DTF²² als ein Anwendungsprofil von ISO 8601 für Datumsangaben** sowie Verwendung von **UNICODE²³ für den Zeichensatz**. Da noch nicht alle deutschen Bibliothekssysteme UNICODE anwenden können, wird der Einsatz von UNICODE für die Zukunft empfohlen.

Auf der lokalen Ebene kann die einzelne Virtuelle Fachbibliothek auch andere Standards, zum Beispiel andere Formen der Datumsangaben, verwenden. Für das gemeinsame Portal muss dann ein entsprechendes Mapping zum empfohlenen Standard erfolgen.

3. Metadatenelemente des VLib Application Profiles

In der vorliegenden Zusammenfassung kann nur ein Ausschnitt mit den wichtigsten Merkmalen der Definition für jedes Element gezeigt werden. Ausführliche Beschreibungen finden sich in den „Empfehlungen für ein Metadaten-Kernset für die Virtuelle Fachbibliothek²⁴“ sowie in der „Kurzfassung“ der Empfehlungen²⁵.

TITLE

Das Metadatenelement *Title* (dc1.1) ist **nicht wiederholbar** und **obligatorisch 1**. Zusätzlich kann das DC Refinement *Title.Alternative* verwendet werden (dcterms, **wiederholbar, empfohlen**). Ein Language-Attribut wird zu beiden Elementen empfohlen, ebenso ein Attribut „Source of content“, in dem angegeben werden kann, ob der Titel vom Autor, vom Bearbeiter oder aus einer dritten Quelle stammt.

Es wird empfohlen, in *Title* den Haupttitel der Ressource anzugeben. Alle anderen Formen sollten unter *Title.Alternative* eingetragen werden.

21 *Media Types* <<http://www.isi.edu/in-notes/iana/assignments/media-types/media-types>>

22 *W3C: Date and Time Formats* <<http://www.w3.org/TR/NOTE-datetime>>

23 *UNICODE homepage* <<http://www.unicode.org/>>

24 Die ausführliche Fassung ist noch nicht öffentlich zugänglich

25 Empfehlungen zur Definition eines Metadaten-Core-Sets für die verteilte Suche über die Virtuellen Fachbibliotheken – Kurzfassung (Version 1.0, 2001-10-05)
<http://www2.sub.uni-goettingen.de/metacore/empfehlungen/core_set.pdf> oder
<http://www.ddb.de/professionell/rtf/vlib_appl_prof.rtf>

Die Einführung eines weiteren Refinements zu Title (Title.Main) wird zur Zeit noch diskutiert; Die Deutsche Bibliothek hat dazu einen Vorschlag in die Diskussionen in der DC-Library Group eingebracht.²⁶

SUBJECT

Das **wiederholbare** Element *Subject* ist insgesamt **obligatorisch 1** und muss mit einem Encoding Scheme (dcterms und/oder VMESq) verwendet werden. Als Encoding Schemes können die in dcterms genannten Thesauri und Klassifikationen (LCSH, MeSH, DDC, LCC, UDC²⁷) verwendet werden oder, da uns diese nicht ausreichend erschienen, weitere spezifische Thesauri und Klassifikationen, die in den VLib Encoding Schemes aufgelistet werden. Dabei sollten vorrangig international verbreitete Sacherschließungssysteme in der jeweils aktuellsten Version verwendet werden. In einem VLib Registry sollen die Versionen mit den entsprechenden Abkürzungen standardisiert verzeichnet werden.

Damit dem Nutzer zukünftig in Der Virtuellen Fachbibliothek eine Navigationsstruktur (cross-browsing) als weiteres Rechercheinstrument angeboten werden kann, wird empfohlen, zusätzlich zu den lokalen Erschließungssystemen die DDC²⁸ bzw. die SWD²⁹ zu verwenden oder sog. Crosskonkordanzen zu diesen Systemen zu erstellen.

Noch nicht eindeutig geklärt ist, welches Metadatenelement (*Subject* oder *Coverage*) sich für die Beschreibung von Geographica besser eignet. In Dublin Core ist eigentlich das Element *Coverage* dafür vorgesehen, inhaltlich gehört es nach Auffassung der VLib Partner eher zu dem Sacherschließungselement *Subject*.

DESCRIPTION

Das Metadatenelement *Description* ist insgesamt **obligatorisch 2** (nicht obligatorisch 1, da nicht alle Katalogdaten ein Element Beschreibung enthalten) und **wiederholbar**. *Description* kann in der unqualifizierten Version (dc1.1) oder in der qualifizierten Version mit den Dublin Core Refinements *Abstract* und/oder *TableOfContents* vorkommen (beide aus dem Namespace dcterms). Genau wie bei *Title* wird ein Language-Attribut empfohlen, ebenso ein Attribut

26 <http://dublincore.org/groups/libraries/>

27 Zur Definition und ausführlicheren Beschreibung siehe:
<http://dublincore.org/documents/dcmes-qualifiers/#subject>

28 Dewey Decimal Classification <<http://www.oclc.org/dewey/>>

29 Schlagwortnormdatei <<http://www.ddb.de/professionell/swd.htm>>

„Source of content“, in dem angegeben werden kann, ob die Beschreibung vom Autor, vom Bearbeiter oder aus einer dritten Quelle stammt.

Als Eintrag für den Elementwert ist Freitext vorgesehen. Offen ist noch, ob das Scheme „URI“ zugelassen werden kann. Dies wäre sinnvoll, sofern eine Einbeziehung verlinkter Ressourcen in das Retrieval möglich ist.

Ob ein weiteres Refinement (*Review*, dann als VMESq Namespace definiert) in das VLib Application Profile integriert wird, ist noch nicht entschieden.

IDENTIFIER

Das Element *Identifier* ist **obligatorisch 1** und **wiederholbar**. Es muss mit dem Dublin Core Value Encoding Scheme *URI* (dcterms) oder einem der VLib Encoding Schemes (z.B. *ISSN*, *ISBN*, *DOI* etc., VMESq) verbunden sein und darf somit nicht in unqualifizierter Version benutzt werden. Als Refinements können *Archive* und/oder *Mirror* (beides VMESq Namespace) benutzt werden.

Insgesamt sollte immer ein eindeutiger *Identifier* für die Ressource angegeben werden, auch wenn es sich um eine Datenbanknummer oder Verbundnummer handelt. Diese Encoding Schemes müssen noch an META-LIB gemeldet werden, damit sie im VLib Application Profile aufgelistet werden können (VMESq).

Einige Praxisempfehlungen für die Vergabe einer URI bei Netzpublikationen sind in den Kurzepfehlungen formuliert worden.

CREATOR, PUBLISHER, CONTRIBUTOR

Die drei Metadatenelemente *Creator*, *Publisher* und *Contributor* (dc1.1) werden aufgrund ihrer engen Verbundenheit hier gemeinsam erörtert. In DCMI befinden sich diese Elemente seit einiger Zeit in intensiver Diskussion, zum Beispiel in der DC-Arbeitsgruppe *agents*³⁰ sowie im Rahmen der Entwicklung des Library Application Profile³¹. Es wird überlegt, ob sie zu einem Metadatenelement zusammengefasst werden können. Attribute und Rollen³² sollen dann jeweils anzeigen, um welche Art von „Agent“ es sich handelt.

Im Empfehlungspapier sind die Elemente *Creator* und *Publisher* **obligatorisch 2**; das Element *Contributor* **empfohlen**. Alle drei Elemente sind **wiederholbar**.

30 <<http://dublincore.org/groups/agents/>>

31 *DC-Library Application Profile (DC-Lib)* <<http://www.dublincore.org/documents/2001/10/12/library-application-profile/>>

32 *Im DC Library Application Profile* wird die *MARC Relator Codes-Liste* <<http://www.loc.gov/marc/relators/re0001.html>> vorgeschlagen.

Die Anwendung der Personennamendatei (PND)³³ und der Gemeinsamen Körperschaftsdatei (GKD)³⁴ als encoding schemes ist zunächst eine in die Zukunft gerichtete Empfehlung Der Deutschen Bibliothek. Voraussetzung dafür ist allerdings nach Meinung der VLib Partner die Entwicklung von Autorentools, die eine arbeitssparende Übernahme der Angaben in die Metadaten und außerdem die gemeinsame Nutzung von anderen, internationalen Normdateien, insbesondere dem LoCNAF, ermöglichen.

Im Empfehlungspapier ist zum jetzigen Stand von der Festlegung weiterer Refinements und Encoding Schemes abgesehen worden. Um kompatibel mit der DC-Entwicklung zu gehen, werden die derzeit laufenden Diskussionen abgewartet. META-LIB beteiligt sich hier und wird die Ergebnisse den Virtuellen Fachbibliotheken vorstellen.

TYPE

Für die Typisierung von Ressourcen wird das Element *Type* in Verbindung mit dem encoding scheme *DCMI Type Vocabulary*³⁵ als **obligatorisch 2** verwendet (dcterms). Es ist **wiederholbar**, falls mehrere Wertangaben zur Charakterisierung der Ressource notwendig sind. Neben dem von DC vorgegebenen Standard können auch eigene Type-Listen von den Virtuellen Fachbibliotheken eingebracht werden (VMESq). Um diese auch für andere nachnutzbar zu machen, sollten sie an ein Registry gemeldet und dort verwaltet werden.

Das von DC vorgelegte Vokabular bietet nur einen Minimalstandard für die Beschreibung der vorliegenden Ressource, der für die Nutzung vielleicht als nicht ausreichend betrachtet wird. Für eine Weiterentwicklung müssen auch andere Listen und Ansätze (z.B. die Arbeitsergebnisse der Arbeitsgruppe Codes der Konferenz für Regelwerksfragen³⁶) wie auch die Gestaltung des Retrievals und der Benutzeroberfläche im Rahmen der übergreifenden Suche über die Virtuellen Fachbibliotheken berücksichtigt werden.

DATE

Das Metadatenelement *Date* ist insgesamt **obligatorisch 2**, muss also, wenn möglich, angegeben werden. Es darf nur mit einem Refinement (dcterms oder VMESq) und einem Encoding Scheme verwendet werden, wobei als Schemes

33 *Personennamendatei* <<http://www.ddb.de/professionell/pnd.htm>>

34 *Gemeinsame Körperschaftsdatei* <<http://www.ddb.de/professionell/gkd.htm>>

35 *DCMI Type Vocabulary* <<http://dublincore.org/documents/dcmi-type-vocabulary/>>

36 <<ftp://ftp.ddb.de/pub/ag-codes>>

W3C-DTF und *DCMI Period* (dcterms) vorgeschlagen werden. Das Element an sich ist **wiederholbar**, die einzelnen Qualifier jedoch sind **nicht wiederholbar**.

Neben den in Dublin Core vorgegebenen Refinements *Date.Created* und *Date.Issued* (dcterms) werden drei neue Qualifier eingeführt: *Date.Submitted* und *Date.Accepted* (VMESq) eignen sich besonders für die Beschreibung von Hochschulschriften.³⁷ Mit *Date.Metadata.Modified* (VMESq) steht ein administratives Element zur Verfügung, um das Datum der letzten Änderung des Metadatensatzes zu verzeichnen.

Über einige weitere Qualifizierungen wird noch diskutiert. Dazu gehören auch die anderen in DC für *date* vorgeschlagenen Qualifier: *valid*, *available*, *modified* sowie die Möglichkeit den Erscheinungsverlauf einer Ressource anzugeben.

FORMAT

Das **wiederholbare** Element *Format* ist nur mit den Refinements *Extent* (**optional**, dcterms) und *Medium* (**empfohlen**, dcterms) zulässig, d.h. es darf nicht ohne Qualifizierung benutzt werden. Als Encoding Scheme für *Extent* wird bei Netzpublikationen die *Internet Media Types* (IMT³⁸) (dcterms) empfohlen.

SOURCE

Das Element *Source* ist **optional** und **wiederholbar** (dc1.1 oder dcterms). Es kann mit dem DC Encoding Scheme „URI“ versehen werden. Dies sollte zum Beispiel bei der Beschreibung von digitalisierten Ressourcen angewendet werden, um mittels einer URI zu der Digitalisierungsgrundlage zu verlinken. Bei der Praxisanwendung von *Source* ergeben sich häufig Abgrenzungsschwierigkeiten gegenüber dem Element *Relation*, unter das *Source* auch subsumiert werden kann. Für die Weiterentwicklung dieses Elements in den Empfehlungen werden Beispiele zur Anwendung von *Source* innerhalb der Virtuellen Fachbibliotheken benötigt.

37 Diese beiden Qualifier wurden aus METADISS übernommen, vgl.
<<http://deposit.ddb.de/metadiss.htm>>

38 <<http://www.isi.edu/in-notes/iana/assignments/media-types/media-types>>

empfohlen. Hier sollten die Encoding Schemes *DCMI Period* oder *W3C-DTF* (beide aus dem Namespace *dcterms*) benutzt werden.

Die Abgrenzung von *Coverage.Spatial* zu *Subject* ist noch nicht endgültig geklärt (s.o.).

RIGHTS

Das Element *Rights* (dc1.1) ist **obligatorisch 2** und **wiederholbar**. Als optionales Encoding Scheme (VMESq) ist die *URI* vorgesehen. Damit soll immer, wenn möglich, eine Angabe zu Copyright-Bedingungen, Kosten bezüglich der Benutzung der Ressource etc. gemacht werden.

Im Moment liegen in Dublin Core keine Qualifizierungen (Refinements, Encoding Schemes) für das Metadatenelement *Rights* vor. Allerdings werden Angaben zur Rechteverwaltung für elektronische Ressourcen derzeit in verschiedenen Projekten etc. diskutiert⁴⁰. Ergebnisse und Erkenntnisse daraus werden in das VLib Application Profile einfließen.

COUNTRY

Das Metadatenelement *Country* bezeichnet den Ort des Publishers bzw. gibt an, in welchen kulturellen Kontext die Ressource eingebunden ist. Es darf nicht benutzt werden, um den sachlichen Inhalt einer Ressource anzugeben. Der Ort als Gegenstand der inhaltlichen Darstellung wird ausschließlich in den Elementen *Subject* bzw. *Coverage* genannt.

Country ist bisher das einzige Element, das nicht in Dublin Core definiert ist (somit VMES Namespace). Es ist **obligatorisch 2**, **wiederholbar** und soll mit dem Encoding Scheme *ISO 3166-1* (VMESq) verwendet werden.

Wahrscheinlich wird die *ISO 3166-1* Liste um einige Codes erweitert, um z.B. europäische und internationale Zusammenhänge abbilden zu können. Auch

40 Siehe z.B. CARMEN AP 2/5 Metadaten zu Terms und Conditions und zur Archivierung <<http://www.sub.uni-goettingen.de/carmen/>>, Reference Model for an Open Archival Information System (OAIS) <<http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-1.pdf>>, ONIX International Standard <<http://www.editeur.org/>>, W3C Workshop on Digital Rights Management <<http://www.w3.org/2000/12/drm-ws/>>, TECUP (Testbed implementation of the European Copyright User Platform) <<http://www.sub.uni-goettingen.de/gdz/tecup/>>, INDECS metadata modell <<http://www.indecs.org/pdf/model1.pdf>>, DOI <<http://www.doi.org>> bzw. <<http://www.doi.org/news/010418-multimedialP.html>>, The Cedars Project: Rights Metadata <<http://www.leeds.ac.uk/cedars/colman/metadata/metadata/spec.html#div36>>

die Einführung einer Hierarchisierung nach Erdteilen (z.B. für Europe, Asia, Africa etc.) wird zur Zeit diskutiert.

4. Ausblick

Das vorgelegte VLib Application Profile stellt den jetzigen Stand dar, bei dem noch offene Fragen geklärt bzw. noch nicht abgeschlossene Entwicklungen berücksichtigt und eingebracht werden müssen. Neben den schon bei den einzelnen Elementen angesprochenen offenen Punkten haben sich bei den Diskussionen von META-LIB mit den Virtuellen Fachbibliotheken noch weitere Fragestellungen ergeben, die hier auszugsweise kurz vorgestellt werden sollen.

Als **Markup-Sprache der Metadaten** wird in den Virtuellen Fachbibliotheken derzeit größtenteils HTML verwendet. HTML weist aber Schwächen bei der Darstellung komplexerer Metadatenstrukturen auf. Daher wird der Übergang von HTML zu einer Encoding-Syntax, die mehrdimensionale Untergliederungen und Schachtelungen zulässt und dabei auf die Verwendung unterschiedlicher Namespaces zurückgreift, sehr empfohlen. Allerdings kann zum jetzigen Zeitpunkt noch keine Empfehlung abgegeben werden, ob in Zukunft die vereinbarten Metadatenelemente in XML⁴¹, RDF⁴² oder in Weiterentwicklungen wie XML Schema⁴³ bzw. RDF Schema⁴⁴ dargestellt werden sollten.

Von entscheidender Bedeutung für das Metadaten-Kernset wird der Aufbau der **Suchmaschine der Virtuellen Fachbibliothek** sein. Die Entscheidung, ob für die Suchmaschine eine zentrale oder eine verteilte Datenhaltung geplant ist, hat Auswirkungen auf die Definition weiterer **administrativer Metadaten**. Es werden auch Festlegungen hinsichtlich der Organisation des Retrievals benötigt: Welche Metadaten sollen bei der Suche (simple, advanced) und beim Browsen genutzt werden; welche Metadatenelemente sollen beim Display (Kurzanzeige / Vollanzeige) der Ergebnisse dem Benutzer angezeigt werden.

Weitere Anforderungen an das Metadaten-Kernset könnten sich im Rahmen der **Archivierung** von Netzpublikationen ergeben. Falls eine Archivierung von Netzpublikationen seitens der einzelnen Virtuellen Fachbibliothek oder durch andere Stellen erfolgt, wären Festlegungen zu **Metadaten zur Archivierung**

41 W3C: *Extensible Markup Language (XML)* <<http://www.w3.org/XML/>>

42 W3C: *Resource Description Framework (RDF)* <<http://www.w3.org/RDF/>>

43 W3C: *XML Schema* <<http://www.w3.org/XML/Schema>>

44 W3C: *Resource Description Framework (RDF) Schema Specification 1.0* <<http://www.w3.org/TR/2000/CR-rdf-schema-20000327/>>

wie auch Metadaten zu dem Bereich „**Rights Management**“ und „**Terms and Condition**“ notwendig.

Generell wird zur Zeit in der internationalen Metadaten-Gemeinschaft viel an der Entwicklung spezifischer Application Profiles gearbeitet. Ein Beispiel für ein solches Application Profile ist das der Arbeitsgruppe DC Education, das zwar hauptsächlich auf Dublin Core Elementen basiert, darüber hinaus aber auch Elemente aus einem anderen Namespace, dem Instructional Management System (Learning Object Metadata, LOM) verwendet. Außerdem werden zusätzliche, domain-spezifische Elemente hinzugefügt, die in einem eigenen Dublin Core Education Namespace definiert werden.⁴⁵ Ein weiteres Beispiel ist das DC-Government Application Profile, dessen Entwicklung sich jedoch noch in der Anfangsphase befindet.⁴⁶ Für die Bibliothekswelt am interessantesten ist das DC-Library Application Profile.

Auch das EU-Projekt Renardus hat für die Definition eines Metadatenkernsets ein spezielles Application Profile entwickelt. Mit Hilfe der dort definierten Metadatenelemente wird eine gemeinsame und gleichzeitige Suche und ein Navigieren entlang der DDC über heterogene und verteilt in Europa vorkommende Subject Gateways ermöglicht.

Um den Austausch und die Nachnutzung von Metadatenelementen, die in Namespaces fest definiert sind, zu ermöglichen, werden diese Informationen an zentralen Stellen, in sog. Registries, gesammelt. Dort kann ein Anwender sich über das Vorhandensein und die Definitionen von Elementen informieren und diese nachnutzen. Zur Zeit gibt es eine Reihe von Registries, die allerdings größtenteils im Aufbau sind oder sich gerade auf neue Anforderungen umstellen:

- das *Dublin Core Registry*, das zur Zeit in zwei unterschiedlichen Prototypen vorliegt und sich auf die in Dublin Core definierten Elemente, Qualifier, kontrollierten Vokabulare etc. beschränkt,⁴⁷
- *MetaForm*, das Mappings von anderen Metadatenformaten zu Dublin Core sammelt und über bestimmte Abfragen die Suche nach speziellen Elementen, Anwendungen etc. erlaubt,⁴⁸

45 *Education Working Group: Draft Proposal* <<http://www.dublincore.org/documents/2000/10/05/education-namespace/>>

46 *DC-Gov Application Profile* <<http://www.dublincore.org/groups/government/profile-200111.shtml>>

47 *The Open Metadata Registry* <<http://wip.dublincore.org:8080/registry/Registry> >

48 *SUB Göttingen - MetaForm* <<http://www2.sub.uni-goettingen.de/metaform/index.html>>

- *Schemas Forum*, ein Registry für alle anwendungsbezogenen Namespaces und Application Profiles. Auch der hier veröffentlichte Prototyp ist noch im Aufbau, eine relationale Datenbank im Hintergrund produziert dynamisch Ergebnislisten. Allerdings ist dieses Registry, im Gegensatz zu dem bei Dublin Core, noch nicht in RDF/XML codiert,⁴⁹
- das *Metadata Observatory* listet mit Schwerpunkt auf Europa alle Projekte auf, die Dublin Core, aber auch anderen Metadatenformate, implementieren.⁵⁰

Über die Definitionen von Namespace und Application Profile wird weiterhin diskutiert. Da ein Application Profile aus mindestens einem Namespace bestehen kann, ist theoretisch jedes Namespace auch ein Application Profile. Dies dürfte z.B. der Grund sein, warum in dem Schemas Forum Registry die beiden Metadatenformate Dublin Core simple und qualified sowohl unter Namespace als auch unter Application Profile definiert sind, die anderen Namespaces dagegen nicht. Hier besteht weiterer Definitionsbedarf, der mit der Erarbeitung von Guidelines für die Entwicklung von Namespaces und Application Profiles durch das Schemas Forum Projekt wahrscheinlich abgedeckt wird. Eine Möglichkeit wäre es z.B., die Definition so einzuschränken, dass ein Application Profile aus mindestens zwei verschiedenen Namespaces oder einer Untermenge eines Namespaces bestehen muss. Bei der, eher unwahrscheinlichen, vollständigen Nachnutzung eines Namespaces mit allen Elementen und ohne die weitere Definition von lokalen Metadatenelementen kann trotzdem ein Application Profile, vielleicht unter Nennung des Projekt- bzw. speziellen Anwendungsnamens, definiert werden. Eine andere Möglichkeit wäre, in diesem Fall kein neues Application Profile zu definieren, sondern in dem Registry bei dem vollständig nachgenutzten Namespace die Anwendungen aufzulisten.

Die Verwendung des hier beschriebenen VLib Application Profiles ermöglicht es dem Nutzer, über einen Einstieg in den verschiedenen Virtuellen Fachbibliotheken zu recherchieren. Das Angebot der bisher 14 Partner wird in Zukunft durch die Beteiligung weiterer Fachbibliotheken noch umfangreicher werden. Zukünftige VLib Partner können sich über das VLib Application Profile über wichtige Metadatenelemente und deren Definition, Anwendung, „best practice“ Empfehlungen etc. informieren und diese nachnutzen.

49 *SCHEMAS Registry* <<http://www.schemas-forum.org/registry/>>

50 *MMI-DC Metadata Observatory* <<http://www.sub.uni-goettingen.de/ssgfi/observatory/>>